



SEQUENCE LISTING

<110> URQUIMA, S.A.

<120> Promotor and constructions for expression of recombinant proteins in filamentous fungi

<130> Thaumatin 2

<140> PCT/EP 99/02243

<141> 1999-04-01

<150> ES 9800699

<151> 1998-04-02

<160> 13

<170> PatentIn Ver. 2.1

<210> 1

<211> 2570

<212> DNA

<213> Aspergillus awamorii

<220>

<221> intron

<222> (785)..(850)

<220>

<221> intron

<222> (1414)..(1471)

<220>

<221> CDS

<222> (741)..(785)

<220>

<221> CDS

<222> (850)..(1413)

<220>

<221> CDS

<222> (1472)..(2242)

<400> 1

tctagattgc gacggcgtat tgcttacccct tagtaggact ccctaattgga ttccgagcaa 60

aaaaagactg tttggcgtgt accaatggct catagtacca gcaagagaag aattttctct 120

ctcgcttcga gaaagcaatc aaaaaaaaaat cctatcctac cctaccctac cctaataactt 180

ccattgccac ccgattccctc ccgatagtag agcggggcag tgccatttgg cggcgccggcc 240

agcggattcc cgccgataga taacgggcag attctgtgac ctcaaactat cgactaacag 300

cccgaaacttc ggccggccacc gccaaaccccg ccccgaaagc cggcctcatt tgccgtttgg 360

gcgtgccagg aatgccgcc tgcagcggag actccctagt gtggctgtg ttgcctgtgt 420

cgtctgtgta gtatactagt tactagtcta ctactgtaca gtggatggcc tgaggggggg 480

actttatgtc cgactccggc tggcttcctc cctctatcca ctctaccctc ttccctctct 540
 tctgtctttc tccccgtctc cgccccctccc ctcctcgaaa acataaaatcg gcctttcccc 600
 ctcgcacatct tttcttctt ctcctctcc tttctcttcc ttcttcagac tacttctctt 660
 tctttcatct tttctctata ttccgtttt cctagatacc ccagttaaaa aagttctctc 720
 aatcaatcct ccccttcaga atg tct aac ctt cct cac gag ccc gag ttc gag 773
 Met Ser Asn Leu Pro His Glu Pro Glu Phe Glu
 1 5 10

cag gcc tac aag ggtatgttcc attgcccctc cgaaattgat gatggaaaaa 825
 Gln Ala Tyr Lys
 15

aaattctaac aacatcctct taca gag ctt gcc acc ctt gag aac tcc 876
 Glu Leu Ala Ser Thr Leu Glu Asn Ser
 20

acc ctc ttc cag aag aac ccc gaa tac cgc aag gcc ctt gct gtc 924
 Thr Leu Phe Gln Lys Asn Pro Glu Tyr Arg Lys Ala Leu Ala Val Val
 25 30 35 40

tcc gtc ccc gag cgt gtc atc cag ttc cgt gtc tgg gag gat gat 972
 Ser Val Pro Glu Arg Val Ile Gln Phe Arg Val Val Trp Glu Asp Asp
 45 50 55

gcc ggc aac gtc cag gtc aac cgc ggt ttc cgt gtc cag ttc aac agc 1020
 Ala Gly Asn Val Gln Val Asn Arg Gly Phe Arg Val Gln Phe Asn Ser
 60 65 70

gcc ctc ggt ccc tac aag ggt ggt ctt cgt ttc cac ccc tcc gtc aac 1068
 Ala Leu Gly Pro Tyr Lys Gly Leu Arg Phe His Pro Ser Val Asn
 75 80 85

ttg tcc atc ctc aag ttc ctt ggt ttc gag cag atc ttc aag aat gct 1116
 Leu Ser Ile Leu Lys Phe Leu Gly Phe Glu Gln Ile Phe Lys Asn Ala
 90 95 100

ctc act ggc ctg aac atg ggt ggt aag ggt ggt tcc gac ttc gac 1164
 Leu Thr Gly Leu Asn Met Gly Gly Lys Gly Ser Asp Phe Asp
 105 110 115 120

ccc aag ggc aag tcc gac aac gag atc cgt cgc ttc tgt gtt tcc ttc 1212
 Pro Lys Gly Lys Ser Asp Asn Glu Ile Arg Arg Phe Cys Val Ser Phe
 125 130 135

atg acc gag ctc tgc aag cac atc ggt gcc gac act gat gtt ccc gct 1260
 Met Thr Glu Leu Cys Lys His Ile Gly Ala Asp Thr Asp Val Pro Ala
 140 145 150

ggt gac atc ggt gtc acc ggt cgt gag gtc ggt ttc ctc ttc ggc cag 1308
 Gly Asp Ile Gly Val Thr Gly Arg Glu Val Gly Phe Leu Phe Gly Gln
 155 160 165

tac cgc aag atc cgc aac cag tgg gag ggt gtt ctc acc ggt aag ggt 1356
 Tyr Arg Lys Ile Arg Asn Gln Trp Glu Gly Val Leu Thr Gly Lys Gly
 170 175 180

ggc agc tgg ggt ggt tcc ctc atc cgc cct gag gcc acc ggt tac ggt 1404
 Gly Ser Trp Gly Gly Ser Leu Ile Arg Pro Glu Ala Thr Gly Tyr Gly
 185 190 195 200

gtt gtc tac gtatgtcaat tcctcttctt atgattatct atgtataaca 1453
 Val Val Tyr

gcgactaacg cgtaacag tac gtc gag cac atg att gct cac gcc acc aac 1504
 Tyr Val Glu His Met Ile Ala His Ala Thr Asn
 205 210

ggc cag gag tcc ttc aag ggc aag cgc gtt gcc atc tcc ggt tcc ggt 1552
 Gly Gln Glu Ser Phe Lys Gly Lys Arg Val Ala Ile Ser Gly Ser Gly
 215 220 225 230

aac gtt gcc cag tac gcc gtc aag gtc att gag ctc ggc ggt tcc 1600
 Asn Val Ala Gln Tyr Ala Ala Leu Lys Val Ile Glu Leu Gly Ser
 235 240 245

gtc gtc tcc ctg agc gac acg cag ggc tcc ctc atc atc aac ggc gag 1648
 Val Val Ser Leu Ser Asp Thr Gln Gly Ser Leu Ile Ile Asn Gly Glu
 250 255 260

ggt agc ttc acc ccc gag gag atc gag ctc atc gct cag acc aag gtc 1696
 Gly Ser Phe Thr Pro Glu Glu Ile Glu Leu Ile Ala Gln Thr Lys Val
 265 270 275

gag cgc aac gag ctc gcc agc atc gtc ggt gct gct ccc ttc agc gac 1744
 Glu Arg Asn Glu Leu Ala Ser Ile Val Gly Ala Ala Pro Phe Ser Asp
 280 285 290

gcc aac aag ttc aag tac att gct ggt gcc cgc ccc tgg gtt cac gtc 1792
 Ala Asn Lys Phe Lys Tyr Ile Ala Gly Ala Arg Pro Trp Val His Val
 295 300 305 310

ggc aag gtc gac gtc gct ctc ccc tcc gct acc cag aac gaa gtt tcc 1840
 Gly Lys Val Asp Val Ala Leu Pro Ser Ala Thr Gln Asn Glu Val Ser
 315 320 325

ggc gag gag gcc cag gtc ctc atc aac gct ggc tgc aag ttc atc gcc 1888
 Gly Glu Glu Ala Gln Val Leu Ile Asn Ala Gly Cys Lys Phe Ile Ala
 330 335 340

gag ggt tcc aac atg ggt tgc acc cag gag gcc atc gac acc ttc gag 1936
 Glu Gly Ser Asn Met Gly Cys Thr Gln Glu Ala Ile Asp Thr Phe Glu
 345 350 355

gcc cac cgt acc gcc aac gct ggc gcg gct gcc atc tgg tac gcc ccc 1984
 Ala His Arg Thr Ala Asn Ala Gly Ala Ala Ala Ile Trp Tyr Ala Pro
 360 365 370

ggt aag gcc gcc aac gcc ggt ggt gtc gct gtc tcc ggt ctg gag atg 2032
 Gly Lys Ala Ala Asn Ala Gly Gly Val Ala Val Ser Gly Leu Glu Met
 375 380 385 390

gct cag aac tct gcc cgc ctc agc tgg act tct gag gag gtt gat gcc 2080
 Ala Gln Asn Ser Ala Arg Leu Ser Trp Thr Ser Glu Glu Val Asp Ala
 395 400 405

cgt ctt aag gac atc atg cgc gac tgc ttc aag aac ggt ctt gag act	2128
Arg Leu Lys Asp Ile Met Arg Asp Cys Phe Lys Asn Gly Leu Glu Thr	
410	415
420	
gct cag gag tac gcc acc ccc gct gag ggt gtc ctg cct tcc ctg gtg	2176
Ala Gln Glu Tyr Ala Thr Pro Ala Glu Gly Val Leu Pro Ser Leu Val	
425	430
435	
acc gga tcc aac att gcc ggt ttc acc aag gtg gct gcc gcc atg aag	2224
Thr Gly Ser Asn Ile Ala Gly Phe Thr Lys Val Ala Ala Ala Met Lys	
440	445
450	
gac cag ggt gac tgg tgg taaatgcgga aagccgcaaa ccccccgcggc	2272
Asp Gln Gly Asp Trp Trp	
455	460
ttatgtcatg acgattatgt agtttgatgt tccctttcag cgccggatgga tagaggcgcc	2332
ggtgttttct tgctagttta gatggatgca taatgatatc cttttcttaa tcctcaaatt	2392
cttgttaattt gttgtatcaa tagtagataa tacaactgta gtcaactacc cttgcatttt	2452
cactatttgc agatgcattc atctctattc cgagcacatg cacaaaccca tgggaccgca	2512
gttcactagt acttagcctg ttatcttccc tctatcgcat cttaaacaac tatctaga	2570

<210> 2
<211> 460
<212> PRT
<213> Aspergillus awamorii

Met Ser Asn Leu Pro His Glu Pro Glu Phe Glu Gln Ala Tyr Lys Glu			
1	5	10	15
Leu Ala Ser Thr Leu Glu Asn Ser Thr Leu Phe Gln Lys Asn Pro Glu			
20	25	30	
Tyr Arg Lys Ala Leu Ala Val Val Ser Val Pro Glu Arg Val Ile Gln			
35	40	45	
Phe Arg Val Val Trp Glu Asp Asp Ala Gly Asn Val Gln Val Asn Arg			
50	55	60	
Gly Phe Arg Val Gln Phe Asn Ser Ala Leu Gly Pro Tyr Lys Gly Gly			
65	70	75	80
Leu Arg Phe His Pro Ser Val Asn Leu Ser Ile Leu Lys Phe Leu Gly			
85	90	95	
Phe Glu Gln Ile Phe Lys Asn Ala Leu Thr Gly Leu Asn Met Gly Gly			
100	105	110	
Gly Lys Gly Gly Ser Asp Phe Asp Pro Lys Gly Lys Ser Asp Asn Glu			
115	120	125	
Ile Arg Arg Phe Cys Val Ser Phe Met Thr Glu Leu Cys Lys His Ile			
130	135	140	

Gly Ala Asp Thr Asp Val Pro Ala Gly Asp Ile Gly Val Thr Gly Arg
 145 150 155 160
 Glu Val Gly Phe Leu Phe Gly Gln Tyr Arg Lys Ile Arg Asn Gln Trp
 165 170 175
 Glu Gly Val Leu Thr Gly Lys Gly Ser Trp Gly Gly Ser Leu Ile
 180 185 190
 Arg Pro Glu Ala Thr Gly Tyr Gly Val Val Tyr Tyr Val Glu His Met
 195 200 205
 Ile Ala His Ala Thr Asn Gly Gln Glu Ser Phe Lys Gly Lys Arg Val
 210 215 220
 Ala Ile Ser Gly Ser Gly Asn Val Ala Gln Tyr Ala Ala Leu Lys Val
 225 230 235 240
 Ile Glu Leu Gly Gly Ser Val Val Ser Leu Ser Asp Thr Gln Gly Ser
 245 250 255
 Leu Ile Ile Asn Gly Glu Gly Ser Phe Thr Pro Glu Glu Ile Glu Leu
 260 265 270
 Ile Ala Gln Thr Lys Val Glu Arg Asn Glu Leu Ala Ser Ile Val Gly
 275 280 285
 Ala Ala Pro Phe Ser Asp Ala Asn Lys Phe Lys Tyr Ile Ala Gly Ala
 290 295 300
 Arg Pro Trp Val His Val Gly Lys Val Asp Val Ala Leu Pro Ser Ala
 305 310 315 320
 Thr Gln Asn Glu Val Ser Gly Glu Ala Gln Val Leu Ile Asn Ala
 325 330 335
 Gly Cys Lys Phe Ile Ala Glu Gly Ser Asn Met Gly Cys Thr Gln Glu
 340 345 350
 Ala Ile Asp Thr Phe Glu Ala His Arg Thr Ala Asn Ala Gly Ala Ala
 355 360 365
 Ala Ile Trp Tyr Ala Pro Gly Lys Ala Ala Asn Ala Gly Gly Val Ala
 370 375 380
 Val Ser Gly Leu Glu Met Ala Gln Asn Ser Ala Arg Leu Ser Trp Thr
 385 390 395 400
 Ser Glu Glu Val Asp Ala Arg Leu Lys Asp Ile Met Arg Asp Cys Phe
 405 410 415
 Lys Asn Gly Leu Glu Thr Ala Gln Glu Tyr Ala Thr Pro Ala Glu Gly
 420 425 430
 Val Leu Pro Ser Leu Val Thr Gly Ser Asn Ile Ala Gly Phe Thr Lys
 435 440 445
 Val Ala Ala Ala Met Lys Asp Gln Gly Asp Trp Trp
 450 455 460

```

<210> 3
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      Oligonucleotide ThS1

<400> 3
      cgaatgaaaa ggaaaaggat ggccacccatc gag      33

<210> 4
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide
      ThS2

<400> 4
      ttattaggcg gtggggca      18

<210> 5
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      Oligonucleotide IA

<400> 5
      atgtctaaacc ttccctcac      18

<210> 6
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      Oligonucleotide IB

<400> 6
      accccttacca ccacccat      18

<210> 7
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      Oligonucleotide IIA

```

<400> 7	
cgcttctgtg tttccttc	18
<210> 8	
<211> 18	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence:Oligonucleotide	
IIB	
<400> 8	
gtacttgaac ttgttggc	18
<210> 9	
<211> 28	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence:Oligonucleotide	
gdh1	
<400> 9	
ttttgtcgac ttgcgacggc gtattgct	28
<210> 10	
<211> 28	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence:Oligonucleotide	
gdh2	
<400> 10	
ttttccatgg tctgaagggg aggattga	28
<210> 11	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence:	
Oligonucleotide Pe	
<400> 11	
gggggttcttc tggaagaggg t	21
<210> 12	
<211> 18	
<212> DNA	
<213> Artificial Sequence	

<220>
<223> Description of Artificial Sequence:
Oligonucleotide NTB2b

<400> 12
atgcgtgctg ctactctc

18

<210> 13
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide CTB2b

<400> 13
ctggccgttg ttgatgag

18